

CS 230
Programming Languages

10 / 04 / 2022

Instructor: Michael Eckmann

Today's Topics

- Questions? / Comments?
- First I present my solution to the parser for the grammar you worked on last week
- More Regular Expressions (Regexs)
- Code examples in Java

Regex

- Let's look here for the 4 principles (what takes precedence when attempting to match) that are followed:

<http://www.cs.rit.edu/~afb/20013/plc/perl5/doc/perlretut.html>

0. Match as early in the input text as possible
1. leftmost alternation that can match, matches
2. quantifiers are greedy
3. if multiple greedy quantifiers, leftmost quantifiers in the regex take precedence for greediness

Regex

- `^` `$` - beginning or end of a line (try w/ and w/o MULTILINE)
 - Note: `^` goes at beginning to match at beginning, `$` goes at end to match at end
- `\A` `\Z` – beginning or end of entire string (not lines within a string) --- useful if using MULTILINE flag and want to match the very beginning or very end of String.

- Let's try these

```
inputText = "Skidmore College\n815 N. Broadway\nSaratoga"
```

```
// note: 2 newlines within the String → three "lines" in the String
```

```
regex = "^[0-9]+";
```

```
// doesn't match in general, will match when MULTILINE flag is on
```

What if we want to store the number it matches 815 in a variable?

Regex

- `.` - matches any character except `\n`
- `.` - w/ DOTALL flag, matches any character including `\n`
- examples:
- `inputText = "Skidmore College\n815 N. Broadway\nSaratoga"`
- `regex = ".*"`;
 - `// matches Skidmore College in general`
 - `// matches entire string if DOTALL flag is on`